Hans Journal of Wireless Communications 无线通信, 2015, 5(6), 126-137
Published Online December 2015 in Hans. http://www.hanspub.org/journal/hjwc
http://dx.doi.org/10.12677/hjwc.2015.56018

Hans汉斯

Survey on Security of Network Coding

Yuyang Zhang[1], Yuanyuan Gao[2], Baofeng Yang[1], Qianqian Zhang[1], Yaokun Gu[3]

[1]College of Communication Engineering, PLA University of Science and Technology, Nanjing Jiangsu
[2]Nanjing University of Posts and Telecommunications, Nanjing Jiangsu
[3]Unit 65040, Shenyang Liaoning
Email: yuyangzhang1991@126.com

Received: Nov. 17th, 2015; accepted: Dec. 2nd, 2015; published: Dec. 16th, 2015

Abstract

In this paper, we mainly summarize secure problems of network coding. Firstly, existing malicious attacks are discussed in detailed definition and classification, and then we describe the relationship and difference among them and point out the basic defense thought of all kinds of malicious attacks. Secondly, aiming at all kinds of malicious attacks, we introduce some classic network coding defense schemes and analyze their advantages and disadvantages. Finally, we summarize existing network coding defense schemes and propose some improvement ideas.

Keywords

Network Coding, Security, Entropy Attack, Byzantine Attack, Pollution Attack, Eavesdropping Attack, Universal Attack, Informationism, Cryptology

网络编码的安全性综述

张宇阳[1]，高媛媛[2]，杨保峰[1]，张倩倩[1]，顾耀坤[3]

[1]解放军理工大学通信工程学院，江苏 南京
[2]南京邮电大学，江苏 南京
[3]65040部队，辽宁 沈阳
Email: yuyangzhang1991@126.com

收稿日期：2015年11月17日；录用日期：2015年12月2日；发布日期：2015年12月16日

文章引用: 张宇阳, 高媛媛, 杨保峰, 张倩倩, 顾耀坤. 网络编码的安全性综述[J]. 无线通信, 2015, 5(6): 126-137.
http://dx.doi.org/10.12677/hjwc.2015.56018

摘　要

本文是针对网络编码的安全性问题的综述性文章，首先对现有的针对网络编码系统的恶意攻击进行了细致的定义和分类，总结了各类恶意攻击之间的联系和区别，说明了它们的基本防御思想；其次针对各类恶意攻击，阐述了一些经典的网络编码防御方案，并分析其优缺点；最后对现有的网络编码防御方案进行了总结并提出了改进思路。

关键词

网络编码，安全，熵攻击，拜占庭攻击，污染攻击，窃听攻击，万能攻击，信息论，密码学

1. 引言

网络编码在 2000 年被 Ahlswede 等人[1]首次提出后便凭借其巧妙的思想、广阔的前景，立刻成为业内的研究热点。网络编码最大的优势就是通过大幅提升网络吞吐量来提高有效性，这是传统的存储-转发中继模式所不具备的。并且，网络编码还能提高网络的鲁棒性和安全性，在一些特殊领域，比如军事和商业领域，安全性是至关重要的。虽然中间节点网络编码后的数据不是直接的有用数据，信宿需要进行译码，这在一定程度上提高了安全性，但是这只是指网络编码在防窃听方面具有得天独厚的优势，而从另一角度看，由于中间节点网络编码后各输入链路的数据组合在了一起，致使一条链路的错误可迅速向整个网络蔓延，这不仅浪费了网络资源，而且会导致信宿译码出错，得不到有用数据，所以网络编码系统的安全性对于主动攻击来说又是非常脆弱的。

本文主要的贡献是：一是对现有的恶意攻击进行了细致的分类和定义，并总结了它们之间的区别与联系，能够帮助读者准确理解概念性问题；二是针对每类恶意攻击，重点介绍了现有的经典网络编码防御方案，能够帮助读者掌握其基本防御思想；三是对现有的网络编码防御方案提出了一些改进思路，可供下一步研究参考。

2. 基本概念

恶意攻击的分类标准有很多，但最常用的是以主动攻击和被动攻击进行区分，其次也有以外部攻击和内部攻击进行区分。主动攻击是指恶意节点主动发送伪造的错误数据来干扰正常通信；被动攻击是指恶意节点并不主动产生并发送数据，只是通过窃听来获取有用数据，该类攻击又称之为窃听攻击。内部攻击是指发动攻击的恶意节点原本是网络内部的合法节点，但是被攻击者利用各种手段进行了策反；外部攻击是指发动攻击的恶意节点是网络外部的非法节点。

根据主动攻击和被动攻击的分类标准，我们将恶意攻击大致分为了熵攻击、广义污染攻击和窃听攻击，再根据内部攻击和外部攻击的分类标准，我们又将广义污染攻击细分为拜占庭攻击和污染攻击。因此，现有的对网络编码系统的恶意攻击可较为细致地分为熵攻击、拜占庭攻击、污染攻击和窃听攻击四类[2] [3]，四类恶意攻击的分类见表 1。

熵攻击：熵攻击是以降低网络编码后数据的差异性为手段，来加长信宿的解码时间，降低网络吞吐量，而在信息论中，信息的差异性就是用信息熵来表示，所以命名此类攻击为熵攻击。熵攻击既可以从网络内部的恶意节点发起，又可以从网络外部的恶意节点发起，所以其既可以是内部攻击，又可以是外部攻击，但它一定是主动攻击。熵攻击可以看作是一种特殊的重放攻击，恶意节点把前时隙接收到的或

者非法窃听到的旧数据在新时隙按合法的网络编码规则进行编码后发出(我们将发出的数据称之为非创新数据)，而不对当前时隙发送的新数据进行编码并发送。又因为非创新数据是旧的数据按照合法网络编码规则编码产生的，所以具有很强的隐蔽性。

见图 1 所示，若熵攻击为内部攻击，A 在时隙 t_1 和 t_2 向下游节点 R_1 分别发送了数据 $\tilde{Y}_{t_1}(e_1)$ 和 $\tilde{Y}_{t_2}(e_1)$，R_1 在接收到后立刻向 B 转发，在时隙 t_3，A 向 R_1 发送了新数据 $\tilde{Y}_{t_3}(e_1)$ 但 R_1 没有立刻向 B 转发，而是利用旧数据按照合法网络编码规则产生了非创新数据 $\tilde{Y}_{t_3}(e_3)$ 并向 B 转发，即 $\tilde{Y}_{t_3}(e_3)=a\tilde{Y}_{t_1}(e_1)+b\tilde{Y}_{t_2}(e_1)$ (a 和 b 是随机选取的编码系数)，这样就会导致 B 得不到 $\tilde{Y}_{t_3}(e_1)$ 但系统又难以检测；若为外部攻击，同理，外部攻击者 C 在非法窃听到 A 发送给下游节点 R_1 和 R_2 的数据 $\tilde{Y}'_{t_1}(e_1)$ 和 $\tilde{Y}'_{t_2}(e_2)$ 后，在时隙 t_3 按照合法网络编码规则产生了非创新数据 $\tilde{Y}'_{t_3}(e'_3)$ 并发送给 B，即 $\tilde{Y}'_{t_3}(e'_3)=a'\tilde{Y}'_{t_1}(e_1)+b'\tilde{Y}'_{t_2}(e_2)$ (a' 和 b' 是随机选取的编码系数)，这样就会干扰 B 得到时隙 t_3 的新数据。

拜占庭攻击：拜占庭攻击是从网络内部发起的，所以其是内部攻击，也正因此，内部恶意节点既有能力窃取到网络传输的有用数据，又有能力主动发送伪造的错误数据来污染有用数据，达到干扰信宿译码，降低网络吞吐量的目的。虽然恶意节点既有窃听又有污染的能力，但是其危害性主要源于主动污染，所以我们将其归为主动攻击一类。因为拜占庭攻击是内部攻击，所以其很难察觉且危害巨大。

污染攻击：污染攻击是指网络外部攻击者向网络主动发送伪造的错误数据来污染有用数据，达到类似于拜占庭攻击的目的，所以其既是外部攻击，又是主动攻击。它和拜占庭攻击的区别除了一个是外部攻击，一个是内部攻击外，还在于一个不具有窃听的能力，而一个具有窃听的能力。图 2 描述了污染攻击在随机线性网络编码中的一个实例。

窃听攻击：对窃听攻击的理解就非常简单了，顾名思义，窃听攻击就是指外部攻击者非法地窃听到网络传输的有用数据。对于有线网络来说，外部攻击者是采用搭线窃听的手段，对于无线网络来说，外部攻击者则是利用高频天线等手段进行窃听。也因为窃听攻击者不会主动发起攻击，所以其是被动攻击，也因此具有隐蔽性。并且人们通常说的被动攻击一般就是特指窃听攻击，见图 3。

四类恶意攻击的联系与区别见表 2。

Table 1. Classification of four malicious attacks
表 1. 四类恶意攻击的分类

名称	内部攻击	外部攻击	主动攻击	被动攻击
熵攻击	√	√	√	
拜占庭攻击	√		√	
狭义污染攻击		√	√	
窃听攻击		√		√

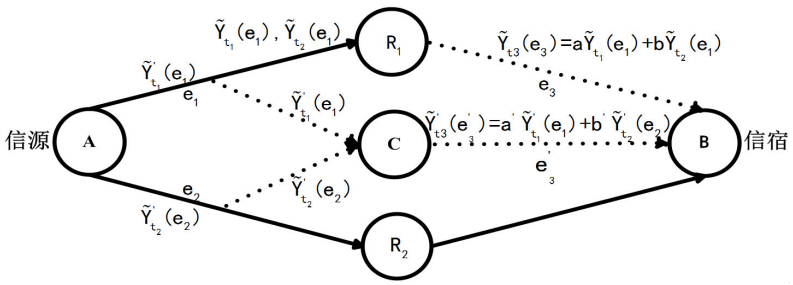

Figure 1. Entropy attack
图 1. 熵攻击

3. 防御方案

目前，针对这四类恶意攻击，现有的网络编码防御方案主要是基于信息论和密码学。

对广义污染攻击来说，现有的网络编码防御方案可大致分为三类：污染纠错、污染检测和污染源定位，其中污染纠错主要是基于信息论，通过构建网络纠错码，而污染检测和污染源定位主要是基于密码学，通过利用同态哈希函数、同态签名、线性子空间签名和认证码等手段。

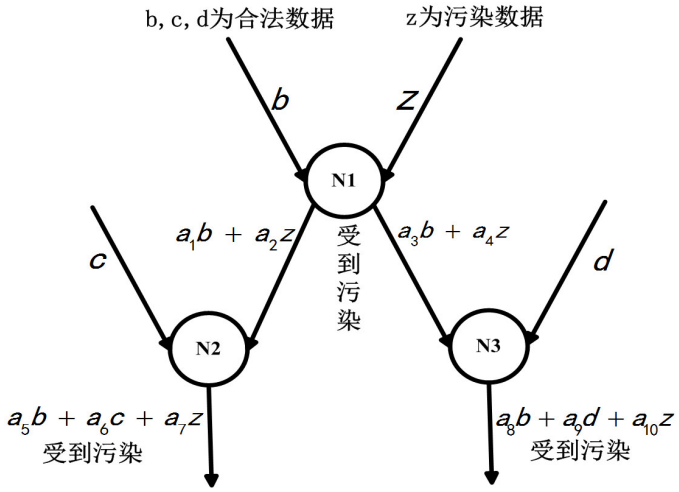

Figure 2. Pollution attack
图 2. 污染攻击

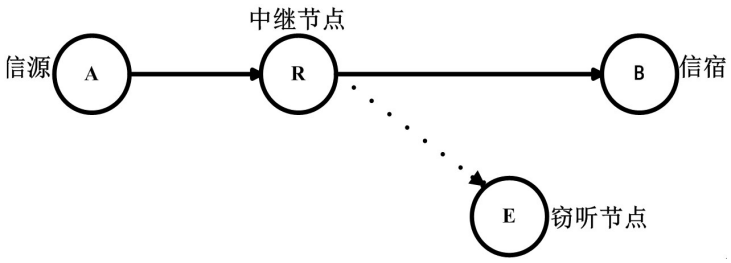

Figure 3. Eavesdropping attack
图 3. 窃听攻击

Table 2. Relationship and difference
表 2. 联系与区别

联系	区别
熵攻击、拜占庭攻击和污染攻击的目的都是主动干扰信宿译码出有用数据，降低网络吞吐量。所以它们都是主动攻击。	窃听攻击的目的是窃听到有用数据，但不主动干扰通信网络。所以它是被动攻击，并且人们常说的被动攻击一般都特指窃听攻击。
拜占庭攻击和污染攻击属于广义污染攻击这一大类，攻击手段都以发送伪造的错误数据来污染有用数据为主。	拜占庭攻击因其为内部攻击，所以内部恶意节点还有窃取网络传输的有用数据这一次要攻击手段。而污染攻击因其为外部攻击，所以不能窃取。
拜占庭攻击和污染攻击发送的是伪造的错误数据，所以发送的数据是非法数据。	熵攻击把前时隙的旧数据在当前时隙按合法的网络编码规则进行编码后发出，所以发送的数据是“合法”数据。
熵攻击因其发送的是“合法”数据，拜占庭攻击因其发动攻击的是“合法”内部节点，窃听攻击因其不主动攻击通信网络，所以都具有隐蔽性。	污染攻击不具有隐蔽性。

对于窃听攻击来说，基于信息论的网络编码防御方案一般是通过限制窃听者的窃听能力来设计网络编码规则，使窃听者译码不出有用数据，从而达到弱安全。而基于密码学的网络编码防御方案一般是对信源数据或编码系数进行加密，使窃听者不能解出有用数据，从而达到弱安全。

3.1. 熵攻击

熵攻击的基本防御思想是通过各种手段来检测接收到的数据是否为非创新数据，若是则舍弃，若不是则进行译码得到有用数据。

Gkantsidis 等人[4]首次提出了熵攻击的概念，并设计了一种防御熵攻击的方案，即检测已编码数据间的线性相关性来判断哪个已编码数据为非创新数据，然后舍弃。但该防御方案的缺点是在一个很大的有限域内进行，所以其计算量很大，耗时很长，大幅降低了网络吞吐量。

Jiang 等人[5]基于自适应概率子集的线性相关性检测算法(S-PSLD)，设计了一种高效的非创新数据过滤方案来防御熵攻击，此方案改进了文献[4]提出的防御方案，大幅降低了计算量。

缓冲监视(BM) [6]的防御思想见图 4 所示，X 是被怀疑需要监视的节点，W 是监视节点，A，B，C 是 X 的上游邻节点。实线表示用于数据传输的无线信道，虚线表示监视节点接收数据的信道。由于无线通信的广播性，W 会接收到 A，B，C 三节点发给 X 的数据和 X 发给下游节点的数据，若 W 发现 X 发出的数据是非创新数据的话，就可以判定 X 为恶意节点。

3.2. 拜占庭攻击

针对拜占庭攻击的网络编码防御方案主要分为基于信息论和基于密码学两类。

3.2.1. 基于信息论的防御方案

Cai等人[7]将经典纠错码的思想引入到网络编码系统中，提出了网络纠错编码这一概念，即在某一时刻通信网络中发生符号错误的链路数没有超出纠错能力范围时，信宿可以通过译码将错误进行纠正。

文献[8] [9]给出了当信道存在噪声时网络纠错编码的具体编译码算法。

Jaggi等人[10]设计了一种分布式随机网络编码防御方案对拜占庭攻击进行纠错。根据恶意节点攻击能力的不同，作者提出了三种算法，但三种算法的基本思想都是将恶意节点视为第二信源。因为信宿接收到的数据是来自信源和恶意节点数据的线性组合，所以只要信宿接收到足够多线性独立的编码数据，就可以译码出来自信源和来自恶意节点的数据，并根据一些限制条件判断出来哪些是来自信源的有用数据，哪些是来自恶意节点的污染数据。

基于信息论的防御方案的纠错或检错效率往往比较高，但是普遍都存在一些局限性：一是都是以攻

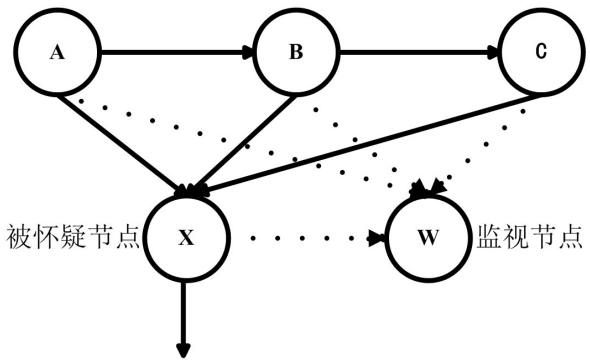

Figure 4. Defense of BM
图 4. BM 防御

击者的攻击能力受限为前提；二是攻击者的攻击能力越强，需要用来纠检错的冗余数据就越多，而冗余数据越多会导致编码率越低，消耗的网络资源越多；三是这些方案只能使信宿被动地对接收到的数据进行纠错或检测，而不能主动地定位、丢弃恶意节点，这样就加长了信宿的译码时间，降低了网络吞吐量。

3.2.2. 基于密码学的防御方案

两不相邻的监视方案[11]假设信源和信宿是可信的，而其他任何中间节点都可能是恶意节点，但只有相邻节点可以共谋。见图5所示是两个X型的网络拓扑结构，从信源到信宿的数据流有两条：S_1到T_1，S_2到T_2。C_1和C_2是进行网络编码的中间节点。每条数据流被两个监视节点监视，因为两个监视节点不相邻，所以它们不能共谋。

以 S_1 到 T_1 这一数据流为例，R_1 是被怀疑节点，S_1 和 R_2 是两个不相邻监视节点，所以 R_1 和 R_2，S_1 和 R_2 之间都不能交流。在第一时隙，S_1 将 a_{t1} 传输至 R_1 和 C_1。在第二时隙，R_1 将 a_{t2} 传输至 T_1 和 C_2。在第三时隙，R_2 从 C_1 接收到网络编码后的数据($a_{t1}+b_{t1}$)，此时 R_2 利用第一时隙从 S_2 得到的 b_{t1} 即可解码得到 a_{t1}。在第四时隙，R_2 从 C_2 接收到网络编码后的数据($a_{t2}+b_{t2}$)，此时 R_2 利用第二时隙它传输的 b_{t2} 即可解码得到 a_{t2}。接下来进行判断，若 $a_{t1}=a_{t2}$，那么 R_1 就不是恶意节点，反之则是。

3.3. 污染攻击

由于污染攻击和拜占庭攻击同属于广义污染攻击，所以针对污染攻击的网络编码防御方案也主要分为基于信息论和基于密码学两类。

3.3.1. 基于信息论的防御方案

防御拜占庭攻击的基于信息论的方案对于污染攻击来说同样适用。

3.3.2. 基于密码学的防御方案

利用同态哈希函数的防御方案：Li 等人[12]利用同态哈希函数来检测污染攻击。在此方案中，见图 6

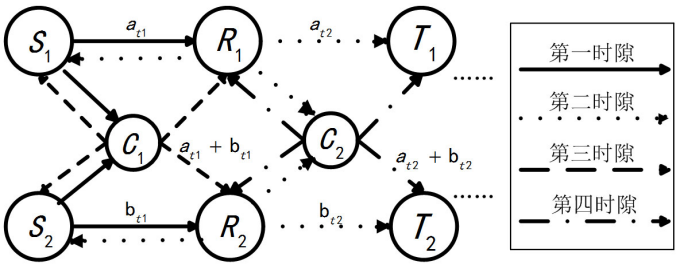

Figure 5. Two X common topologies
图 5. 两个 X 型网络拓扑结构

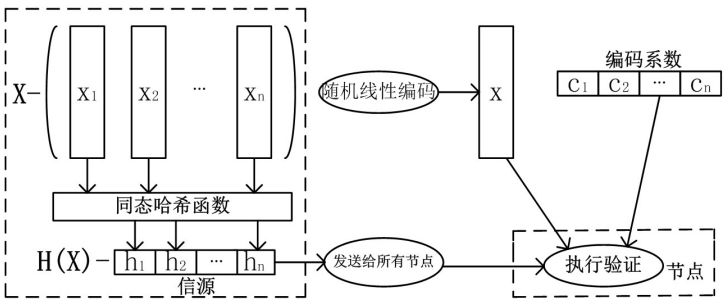

Figure 6. Homomorphic Hash and verification
图 6. 利用同态哈希函数的防御方案

所示，数据 x 被分为 $x_1, x_2, \cdots, x_n$ 这 n 块数据块，每块 x_i 又被进一步分为 $x_{i,1}, x_{i,2}, \cdots, x_{i,m}$ 这 m 块子数据块。$x_{i,j} \in Z_p^* \left(j = 1, 2, \cdots, m\right)$。将哈希函数 $H(x)$应用于 x_i 上得到相应哈希值的 h_i，$h_i = \prod_{j=1}^{m} g_j^{x_{i,j}} \bmod p$，$g_j \in Z_p^* \left(j = 1, 2, \cdots, m\right)$，并且哈希函数 $H(x)$具有同态性质即 $H\left(x_i\right) H\left(x_j\right) = H\left(x_i + x_j\right)$。这 n 个哈希值 $h_i \left(1, 2, \cdots, n\right)$ 被提前发至各个节点。因为网络编码后的 x 应该是 $x_i \left(1, 2, \cdots, n\right)$ 的线性组合，编码系数为 $C = \left(c_1, c_2, \cdots, c_n\right)$，即 $x = c_1 \cdot x_1 + c_2 \cdot x_2 + \cdots + c_n \cdot x_n$，所以根据哈希函数的同态性，如果中间节点或信宿接收到的数据的哈希值等于 $\prod_{i=1}^{n} h_i^{c_i} \bmod p \left(i = 1, 2, \cdots, n\right)$，那么消息没被污染，反之则已被污染。

利用线性子空间签名的防御方案：基于正交向量检测污染攻击的网络编码防御方案[13]的基本思想是计算出子空间的正交向量并将其发送给中间节点和信宿，这些节点可以利用正交向量检测接收到的数据向量是否被污染。比如，信源发送h个数据，每个数据是N维向量，那么信源发送的消息矩阵就是：

$$\begin{bmatrix} x_1 \\ \vdots \\ x_h \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 & x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & x_{h,1} & x_{h,2} & \cdots & x_{h,N} \end{bmatrix} \tag{1}$$

消息矩阵中 $x_{i,j} \in F_q$，网络中边 e 上传输的数据是x_i的线性组合，记为 $y(e) = \sum_{i=1}^{h} g_i(e) x_i$，$g_i(e)$ 是全局编码向量，所以 $y(e) \in F_q$。因为向量 $x_1, x_2, \cdots, x_h \in F_q^{h+N}$ 且线性独立，所以它们可以构成一子空间 X，由于 $y(e)$是 $x_i \left(i = 1, 2, \cdots, h\right)$ 线性组合，所以 $y(e) \in X$。子空间 X 中的每一个向量都可以视为合法数据，那么求出满足 $x_i \cdot v = 0$ 的子空间 X 的正交向量 $v = \left(v_1, v_2, \cdots, v_{h+N}\right) \in F_q^{h+N}$。将正交向量 v 提前发给网络中的其它节点，若中间节点或信宿接收到的数据向量点乘 v 等于 0 则数据没有被污染，反之则已被污染。并且可以根据上游节点的点乘值来判断恶意节点的具体位置。

3.4. 窃听攻击

针对窃听攻击的网络编码防御方案也主要分为基于信息论和基于密码学两类。

3.4.1. 基于信息论的防御方案

Cai 等人[14]提出了搭线窃听的网络通信模型 GSTW，并构造了在信息论意义下的安全网络编码防御方案，即窃听者无论窃听所给定窃听集内的哪个窃听子集都无法恢复出原始数据。随后，他们进一步提出了 r-安全网络编码防御方案[15] [16]，该方案也需限制窃听者的窃听能力，即窃听信道数小于加入的随机数个数 r。

Bhattad 等人[17]适当合理地降低了安全条件，提出了弱安全网络编码的概念，只要窃听者无法获得任何有用数据即可。当窃听者不能获得信源发送的任何原始数据时，我们称之为信息论安全网络编码；当窃听者不能获得信源发送的任何有用数据时，我们称之为弱安全网络编码。由于弱安全网络编码放宽了安全条件，但又很好地保护了有用数据，相对于信息论安全网络编码来说，又提高编码的最大多播速率，因此，现在人们研究的防窃听网络编码防御方案基本上都是追求达到弱安全。

文献[18]已知网络的拓扑结构，设计了防窃听的网络编码防御方案。图 7 是窃听者位置已知的简单双向通信防御方案：s 为信源，d 为信宿，e 为窃听节点，假设 e 的窃听能力受限，因为 e 在靠近 s 的 sd 直线上，所以 s 广播给 d 的数据会被 e 窃听到，而 d 广播 s 的数据则不会。因此，根据此网络拓扑结构设计的网络编码防御方案是：第一步，d 产生一个随机数 k 并发送给 s，第二步，s 利用接收到的 k 对有用数据 x 进行网络编码如 c = x ⊕ k 然后发送给 d。则 d 有 c 和 k 即可解码出 x，而 e 只有 c 而没有 k 则解码不出 x。

图 8 是两个窃听者位置已知但不共谋的防御方案：s，d，a，b 是合法节点，e_1，e_2 是两个位置已知但不共谋的窃听节点，e_1 在 s 到 a 的传输路线上，e_2 在 s 到 b 的传输路线上。根据此网络拓扑结构设计的

网络编码防御方案是：a，b 分别生成随机数 k_1 和 k_2 并发送给 s 和 d，然后 s 将进行网络编码 $c = x \oplus k_1 \oplus k_2$ 并将 c 发送给 a 和 b，最后由 a 或者 b 将 c 发送给 d。d 由于接收到 c，k_1，k_2 就可以解码出 x，而 e_1 由于只窃听到 c 和 k_1，e_2 由于只窃听到 c 和 k_2，所以都不能解码出 x。

图 9 是两个窃听者位置未知但不共谋的防御方案：s，d，r_1，r_2，r_3，r_4 是合法节点，e_1，e_2 是两个位置未知但不共谋的窃听节点，并且窃听节点位置不会与合法节点位置相重。在此，我们假设两个窃听节点位于最好的窃听位置(见图 9 所示分别位于两个方格拓扑结构中央，这样就可以窃听到最大量的数据)。根据此拓扑结构设计的网络编码防御方案是：第一步，r_1，r_2 分别生成随机数 k_1，k_2 并沿发送路径发送，第二步，s 利用 k_1，k_2 进行网络编码 $c = x \oplus k_1 \oplus k_2$ 并沿发送路径发送。各路径传输完毕后可知，d 由于接收到 c，k_1，k_2 所以可以解码出 x，而 e_1 和 e_2 由于分别缺少 k_2 和 k_1 所以解码不出 x。

3.4.2. 基于密码学的防御方案

基于信息论的防御方案以限制窃听者的窃听能力为前提，当它们面对窃听能力更强的窃听者时便不能保证网络的安全性。因此，用密码学手段来设计防窃听的网络编码防御方案受到了广泛的关注。

通过对编码后数据及编码向量进行排列加密操作的防窃听的网络编码方案，称之为 P-Coding [19]。其基本思想见图 10 所示，经过排列加密操作后，编码后数据和其编码向量可以混合和重新排序。由于在随机网络编码中的编码后数据是源数据的组合，要恢复出源数据就需要利用编码向量对其进行解码。而排列加密函数将编码向量的元素和编码后数据的元素随机混合后，窃听者很难定位到哪些是编码向量的

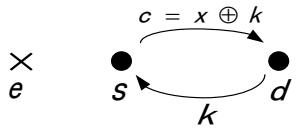

Figure 7. A simple two-way scheme
图 7. 简单双向通信防御方案

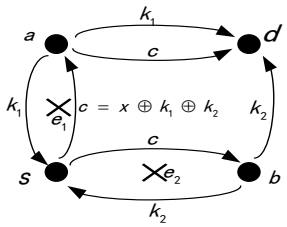

Figure 8. Non-collaborating eavesdroppers of known location
图 8. 两个窃听者位置已知但不共谋防御方

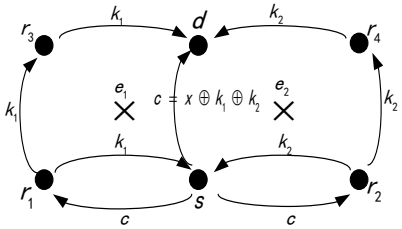

Figure 9. Eavesdroppers of unknown location
图 9. 两个窃听者位置未知但不共谋防御方案

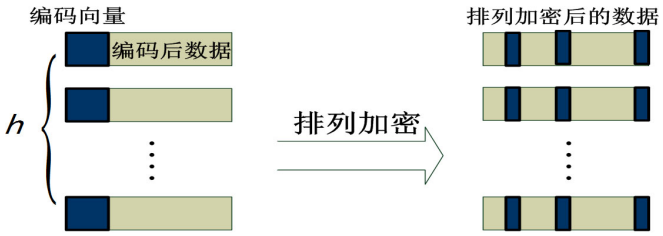

Figure 10. Permutation encryption on coded messages
图 10. 在已编码数据上进行排列加密

元素，哪些是编码后数据的元素，也就无法解码出源数据。

防窃听的 VSWNC 方案[20]要求信源信宿共享一个随机数生成器，其次信源可用随机数 r 在随机数生成器上生成一个范德蒙行列式 P，随后用 P 对信源数据进行预编码处理，最后信宿可通过 r 得到 P 从而能够正确解码。把 VSWNC 方案应用于随机网络编码中，在牺牲少量带宽的情况下能保证以概率 1 达到弱安全。

信源编码：

信源在有限域 M 中选取一个随机数 r，然后用 r 在随机数生成器上生成 $n-1$ 个随机数。$r_1, r_2, \cdots, r_{n-1}$ 这 $n-1$ 个数构成了范德蒙行列式。r 是对窃听者绝对保密的。

$$P = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & r_1 & \cdots & r_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_1^{N-1} & \cdots & r_{n-1}^{N-1} \end{bmatrix} \tag{2}$$

X左乘P可得到 $X' = PX$ 。

$$X' = \begin{bmatrix} x'_{11} & x'_{12} & \cdots & x'_{1M} \\ x'_{21} & x'_{22} & \cdots & x'_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{N1} & x'_{N2} & \cdots & x'_{NM} \end{bmatrix} \tag{3}$$

在 X' 后加入单位冗余得到 Y，其中 $p_1, p_2, \cdots, p_{N-1}$ 为 M 上的随机数。

$$Y = \begin{bmatrix} x'_{11} & x'_{12} & \cdots & x'_{1M} & r \\ x'_{21} & x'_{22} & \cdots & x'_{2M} & p_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x'_{N1} & x'_{N2} & \cdots & x'_{NM} & p_{N-1} \end{bmatrix} \tag{4}$$

经过以上变换，信源就把Y发送出去。Y在网络中进行网络编码，即$Z = FY$。

信宿解码：

信宿收到 Z 后可用公式 $Y = F^{-1}Z$ 获得 Y，然后得到 r 和 X'。信宿再利用 r 访问随机数生成器，来获得 $r_1, r_2, \cdots, r_{N-1}$，进而得到 P。又因为 $X = P^{-1}X'$，所以信宿就可以获得信源数据 X 了。

3.5. 万能攻击

就像我们会想方设法地保护网络的安全性一样，攻击者也会采取各种手段来达成他们的目的。在现实中，攻击者往往会采用多类恶意攻击来同时攻击网络，因此，设计出能同时防御多类恶意攻击的网络编码防御方案是大势所趋，我们称此类防御方案为抗万能攻击方案。

文献[21]利用哈希函数和安全信道对随机单跳网络编码进行改进，采用一个 τ 维随机向量 α 和两个哈希函数 $\theta_1(\cdot)$，$\theta_2(\cdot)$ 来同时防御窃听攻击和广义污染攻击。该方案仅改变了信源和信宿的编解码方式，而未改变中间节点的网络编码方式。

信源编码：

信源产生的原始数据 M 要经过 2 次变换：

$$M=\begin{Bmatrix} m_{11} & m_{12} & \cdots & m_{1k} \\ m_{21} & m_{22} & \cdots & m_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nk} \end{Bmatrix}=\begin{Bmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{Bmatrix} \tag{5}$$

$M_i(i=1,2,\cdots,n)$ 为原始数据向量，其中 $M_1 \in F_q^k$。在有限域 $GF_{(q)}$ 上选取随机向量 α，且服从均匀分布。使用哈希函数 $\theta_1(\cdot): F_q^\tau \to F_q^k$ 使得 $\theta_1(\alpha) \in F_q^k$。我们通过 α 构建新的数据向量：

$$\begin{cases} M_1' = M_1 + \theta_1(\alpha) \\ M_2' = M_2 + \theta_1(\alpha, M_1) \\ \quad \vdots \\ M_n' = M_n + \theta_1(\alpha, M_1, M_2, \cdots, M_{n-1}) \end{cases} \tag{6}$$

经过第一次变换，得到了新的数据矩阵 $M'=\{M_1', M_2', \cdots, M_n'\}^{\mathrm{T}}$。使用哈希函数 $\theta_2(\cdot)$：$F_q^k \to F_q^l$，得到哈希值 $\theta_2(M_i) \in F_q^l\ (i=1,2,\cdots,n)$ 并附加在 $M_i'(i=1,2,\cdots,n)$ 数据向量之后，使之形成新的向量 M_i''。

$$M''=\begin{Bmatrix} M_1' & \theta_2(M_1) \\ M_2' & \theta_2(M_2) \\ \vdots & \vdots \\ M_n' & \theta_2(M_n) \end{Bmatrix}=\begin{Bmatrix} M_1'' \\ M_2'' \\ \vdots \\ M_n'' \end{Bmatrix} \tag{7}$$

经过第二次变换后，信源形成了新的数据向量 $M_i''(i=1,2,\cdots,n)$，每个数据向量由 $k+l$ 位符号组成，前 k 位符号为数据消息 M_i'，后 l 位符号由 M_i 的哈希函数值 $\theta_2(M_i)$ 构成。

信宿译码：

信宿收到一个消息矩阵 D，其由 n 个线性独立的数据组成。

$$D=M''\times T=\begin{Bmatrix} M_1'' \\ M_2'' \\ \vdots \\ M_n'' \end{Bmatrix}\times T \tag{8}$$

T 为 n 个数据系数组成的系数矩阵，求出逆矩阵 T^{-1}，解码出数据矩阵 M''。数据向量 $M_i''(i=1,2,\cdots,n)$ 的前 k 位是数据消息 M_i'，通过安全信道接受到随机向量 α，信宿可以从 M_i' 中解出 M_i。稍后信宿通过哈希函数 $\theta_2(\cdot)$ 计算出 $\theta_2(M_i)$，与收到的矩阵对比，如果值相同，则说明没有受到污染攻击。反之则受到，那么就丢弃此数据矩阵 D。

对于窃听攻击，由于随机向量 α 是由安全信道传输的，攻击者无法窃听到，即使攻击者窃听到 n 个线性独立的数据，解码出 M''，但由于没有得到随机向量 α，因此无法解码出 M，所以该方案防窃听攻击。并且，攻击者虽然知道了哈希函数 $\theta_2(\cdot)$，但是没有解码出 M，因而无法得到正确的哈希值 $\theta_2(M_i)$，所以该方案也防广义污染攻击。然而，信宿只能检测出污染并丢弃整个数据矩阵，而不能进行纠错，这

样就延长了传输时间，降低了网络吞吐量。Xu 等人[22]先是利用稀疏矩阵对信源数据进行矩阵变换达到了防窃听的目的，再利用列表译码法在信宿处进行译码，有效地进行了纠错，也达到了防广义污染攻击的目的，并进一步缩短了传输时间。

在文献[5]的基础上出现了一种增强型的抗万能攻击方案[23]。该方案通过计算信宿接收到的数据的线性相关性来判断其是否为非创新数据，再检测其是否为污染数据，只要有一样是，就将其过滤掉，达到了既防熵攻击又防广义污染攻击的目的。

4. 总结

网络编码从根本上改变了通信网络的中继转发模式，具有巨大的潜力，而安全性作为限制其发展应用的一大制约指标，设计网络编码防御方案已成为一个研究热点。分析总结现有的网络编码防御方案，我们提出了一些改进思路，供读者参考：

1) 网络编码的最大优势是提高通信网络的有效性，而信道编码是提高可靠性，研究怎样将网络编码和信道编码完美结合则可以同时提高通信网络的有效性和可靠性。

2) 在防御广义污染攻击时，可以考虑建立集检错和纠错于一体的网络编码防御方案。

3) 大多数现有的网络编码防御方案只针对一类或两类恶意攻击，或者限制了攻击者的攻击能力。例如，针对拓扑结构设计的 NC 方案来防御窃听者们也是假设窃听者们不共谋罢了。随着攻击技术的快速发展，我们的防御方案也需要在深度上和广度上快速发展，最好能实现深度和广度的完美结合，达到真正意义上的抗万能攻击。

4) 一些防御方案需要特殊的网络拓扑结构或额外的安全信道，如方格网络拓扑结构、两个 X 型网络拓扑结构等。在现实的无线网络系统中，这些拓扑环境都很难满足，所以适用性更强的网络编码防御方案是未来的趋势。

参考文献 (References)

[1] Ahlswede, R., Cai, N. and Yeung, R.W. (2000) Network Information Flow. *IEEE Transactions on Information Theory*, **46**, 1204-1216. http://dx.doi.org/10.1109/18.850663

[2] Yao, S.X., Chen, J., *et al.* (2014) A Survey of Security Network Coding toward Various Attacks. 2014 *IEEE* 13*th International Conference on Trust, Security and Privacy in Computing and Communications*, Beijing, 24-26 September 2014, 252-259. http://dx.doi.org/10.1109/TrustCom.2014.35

[3] He, M., Chen, L., *et al.* (2012) Survey on Secure Transmission of Network Coding in Wireless Networks. 2012 *International Conference on Computer Science and Service System*, Nanjing, 11-13 August 2012, 1216-1219. http://dx.doi.org/10.1109/CSSS.2012.308

[4] Christos, G. and Pablo, R. (2006) Cooperative Security for Network Coding File Distribution. *Infocom*, **3**, 5.

[5] Jiang, Y.-X., Fan, Y.-F., Xue, M., *et al.* (2009) A Self-Adaptive Probabilistic Packet Filter Scheme against Entropy Attacks in Network Coding. *Computer Networks*, **53**, 3089-3101. http://dx.doi.org/10.1016/j.comnet.2009.08.002

[6] Newell, A.J., Curtmola, R. and Nita-Rotaru, C. (2012) Entropy Attacks and Countermeasures in Wireless Network Coding. *Proceedings of the Fifth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, New York, 16-18 April 2012, 185-196. http://dx.doi.org/10.1145/2185448.2185473

[7] Cai, N. and Yeung, R.W. (2002) Network Coding and Error Correction. *IEEE Information Theory Workshop*, Bangalore, 20-25 October 2002, 119-122.

[8] Zhang, Z. (2006) Network Error Correction Coding in Packetized Networks. *IEEE Information Theory Workshop*, Chengdu, 433-437. http://dx.doi.org/10.1109/itw2.2006.323836

[9] Zhang, Z. (2008) Linear Network Error Correction Codes in Packet Networks. *IEEE Transactions on Information Theory*, **54**, 209-218. http://dx.doi.org/10.1109/TIT.2007.909139

[10] Jaggi, S., Langberg, M., Katti, S., Ho, T., Katabi, D. and Medard, M. (2007) Resilient Network Coding in the Presence of Byzantine Adversaries. 26*th IEEE International Conference on Computer Communications*, Anchorage, 6-12 May 2007, 616-624. http://dx.doi.org/10.1109/INFCOM.2007.78

[11] Pandit, V., Jun, J.H. and Agrawal, D.P. (2011) Inherent Security Benefits of Analog Network Coding for the Detection of Byzantine Attacks in Multi-Hop Wireless Networks. *Proceedings of the* 2011 *IEEE* 8*th International Conference on Mobile Ad Hoc and Sensor Systems* (*MASS*), Valencia, 17-21 October 2011, 697-702.

[12] Li, Q., Lui, J.C.S. and Chiu, D.-M. (2012) On the Security and Efficiency of Content Distribution via Network Coding. *IEEE Transactions on Dependable and Secure Computing*, **9**, 211-221.

[13] Bin, D., Zhang, S., Qu, Y., Yang, J. and Wang, F. (2010) Orthogonal Vector Based Network Coding against Pollution Attacks in n-Layer Combination Networks. *Proceedings of the* 5*th International ICST Conference on Communications and Networking in China* (*CHINACOM*), Beijing, 22-25 August 2010, 1-5.

[14] Cai, N. and Yeung, R.W. (2002) Secure Network Coding. *Proceedings of the IEEE International Symposium on Information Theory*, Lausanne, 30 June-5 July 2002, 323.

[15] Cai, N. and Yeung, R.W. (2011) Secure Network Coding on a Wiretap Network. *IEEE Transactions on Information Theory*, **57**, 424-435. http://dx.doi.org/10.1109/TIT.2010.2090197

[16] Cai, N. and Chan, T. (2011) Theory of Secure Network Coding. *Proceedings of the IEEE*, **99**, 421-437. http://dx.doi.org/10.1109/JPROC.2010.2094592

[17] Bhattad, K. and Narayanan, K.R. (2005) Weakly Secure Network Coding. *Proceedings of the First Workshop on Network Coding*, *Theory*, *and Applications*, Riva del Garda, 7 April 2005, 281-285.

[18] Capar, C. and Goeckel, D. (2012) Network Coding for Facilitating Secrecy in Large Wireless Networks. *Proceedings of the* 46*th Annual Conference on Information Sciences and Systems* (*CISS*), Princeton, 21-23 March 2012, 1-6.

[19] Zhang, P., Jiang, Y., Lin, C., Fan, Y. and Shen, X. (2010) P-Coding: Secure Network Coding against Eavesdropping Attacks. *Proceedings of the IEEE INFOCOM*, San Diego, 14-19 March 2010, 1-9.

[20] 武萌, 吴蒙. 防窃听的弱安全网络编码[J]. 计算机技术与发展, 2014, 24(10): 167-169.

[21] 赵佳佳, 任平安. 基于抗窃听和拜占庭攻击的随机网络编码[J]. 计算机科学, 2014, 41(9): 174-177.

[22] 徐光宪, 付晓. 抗万能攻击的安全网络编码[J]. 计算机科学, 2012, 39(8): 88-91.

[23] 杨柳, 钟诚. 一种高效的安全数据包过滤算法[J]. 兰州大学学报(自然科学版), 2012, 48(4): 105-108.